

Academy of Aphasia 2010

Simulating Bilingual Aphasia: A Novel Computational Model

Uli Grasemann^{a,*}, Swathi Kiran^b, Chaleece Sandberg^b, Risto Miikkulainen^a^a *Computer Science, University of Texas at Austin*^b *Speech and Hearing Sciences, Boston University*

Introduction

Current research on bilingual aphasia has only begun to inform us about the optimal rehabilitation for bilingual aphasic patients (Roberts & Kiran, 2007; Edmonds & Kiran, 2006) but the literature is still sparse in terms of interpreting the nature of naming impairments in bilingual aphasia. We have developed a computational model to simulate a bilingual language system in which language representations can vary by age of acquisition and relative proficiency in the two languages. Based on recent theoretical models, a single semantic map is used, with separate maps for the words in the two languages, and six separate sets of connections between the maps. This model is subsequently lesioned at specific sites by varying connection strengths between the semantic and phonological networks.

Method

Currently, this model simulates a Spanish-English bilingual lexicon but in the future can be extended to any language combination. L1 is Spanish, L2 is English, and the AoA and proficiency are continuous variables. The semantic map consists of 300 nouns obtained from our previous studies and comprises of English-Spanish translations. Each noun is encoded based on a set of 260 binary features such as "can lay eggs", "can be slept on", and "can be used as a weapon". The resulting numerical representations of the words were then used to train the semantic map of the lexicon. English and Spanish phonetic encodings of the words were derived automatically from their IPA transcriptions. The three self-organizing maps (semantic, L1 and L2) each had 30x40 neurons, and were trained simultaneously with the associative connections between each pair of maps. All three maps used Gaussian neighborhood functions whose width decreased exponentially (from $\sigma = 7.0$ down to 0.2) over the course of training. Later, age of acquisition for L2 was simulated by delaying the onset of training for the L2 phonetic map and its associative connections. Differences in exposure to L1 and L2 were simulated by exposing the model to more or less phonetic input in each language. Learning rates for all maps and all associative connections were 0.25. Each model was trained for 1000 epochs.

Results

In order to match the model's performance in both English and Spanish to that of a group ($N = 39$) of individual bilingual human speakers with varying AoA and relative proficiency, the training parameters were set up to match the known ages of acquisition and exposure data as closely as possible for each test case. In most cases, the model is able to replicate the performance of human subjects quite closely. We then extended the model to simulate a group of bilingual aphasia patients ($N = 19$), by attempting to replicate the patients' self-reported AoA and pre-stroke performance. Again, the model usually comes close to the target performance in both languages. Cases where the model does not match the patient data can be attributed to a mismatch between

* Corresponding author.

E-mail address: uli.grasemann@gmail.com.

model's performance that is based on an estimate of exposure whereas the actual patient's learning history is more complicated and hence more variable.

References

- Edmonds, L., & Kiran, S. (2006). Effect of semantic naming treatment on crosslinguistic generalization in bilingual aphasia. *Journal of Speech Language and Hearing Research*, 49, 729-748.
- Roberts, P. & Kiran, S. (2007). Assessment and treatment of bilingual aphasia and bilingual anomia. A. Ardila & E. Ramos (Eds.) *Speech and Language Disorders in Bilinguals*. Nova Science Publishers. New York, NY, pp 109-131.

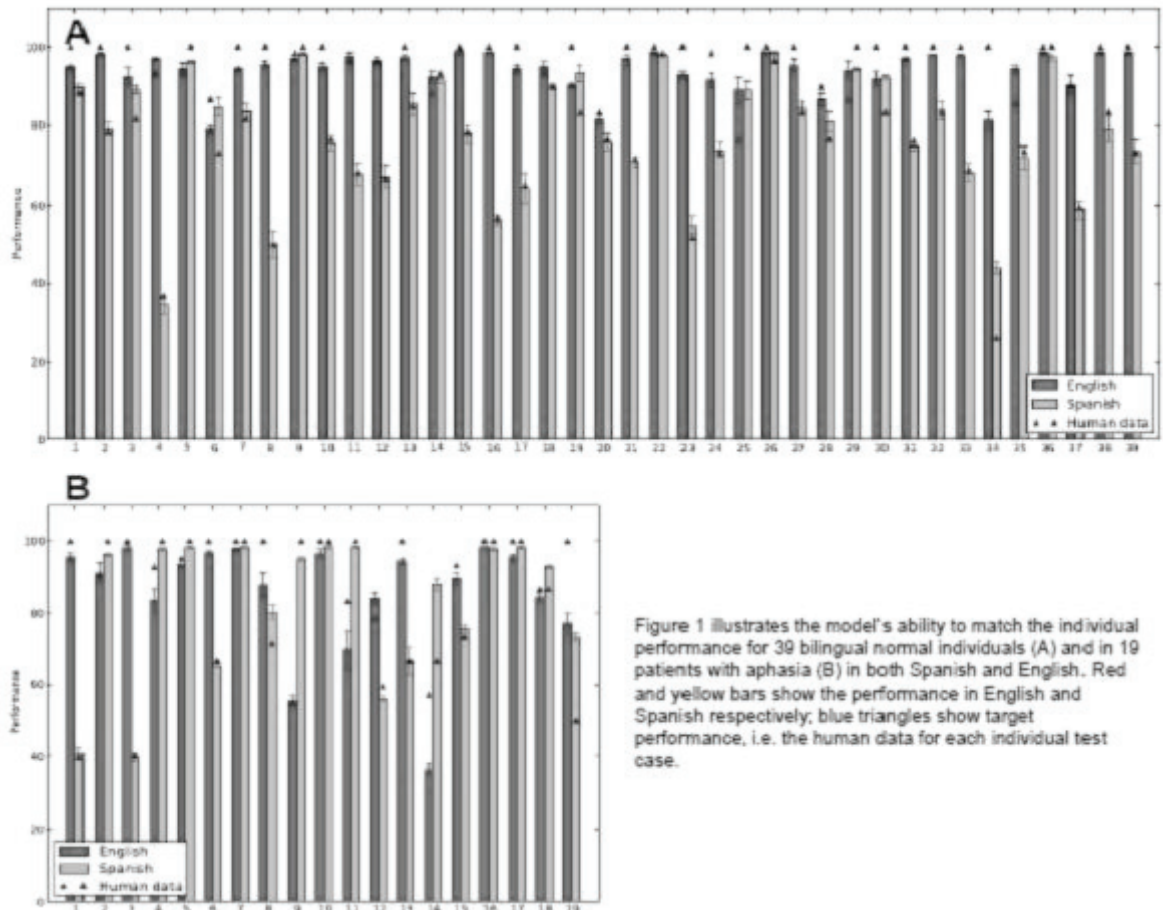


Figure 1 illustrates the model's ability to match the individual performance for 39 bilingual normal individuals (A) and in 19 patients with aphasia (B) in both Spanish and English. Red and yellow bars show the performance in English and Spanish respectively; blue triangles show target performance, i.e. the human data for each individual test case.